

# Generation of polyketide libraries via combinatorial biosynthesis

Chaitan Khosla and Robert J. X. Zawada

Polyketides are a family of structurally complex natural products that include a number of important pharmaceuticals. Motivated by the value of these natural products, there has been much research focused on developing guidelines for engineering polyketide synthases (PKSs) to generate novel polyketides. Recent studies have provided interesting insights into the enzymatic specificity of the polyketide synthesis pathway, and have demonstrated that various PKSs can be genetically manipulated to synthesize 'unnatural' polyketide natural products. In this article, we discuss the synthesis of polyketides and polyketide libraries by combinatorial biosynthesis.

Polyketides are a large family of complex and structurally diverse natural products, and an extremely rich source of bioactive molecules. They include numerous antibiotics, anticancer agents, immunosuppressants, antiparasitic agents, antifungals, cardiovascular agents and veterinary products. Like the related fatty acid synthases (FASs), the polyketide synthases (PKSs) are multi-enzyme assemblies that catalyze repeated decarboxylative condensations between coenzyme A thioesters<sup>1</sup>. Following each condensation, FASs typically catalyze a complete reductive cycle comprising a ketoreduction, dehydration and enoylreduction on the  $\beta$ -keto group of the growing carbon chain; PKSs, however, omit this cycle or curtail it after some, or even all, post-condensation steps. After the carbon chain has grown to a length characteristic of each specific product, it is released from the synthase by thiolysis or acyltransfer. It is the controlled variation in chain length, choice of chain-building units and the differences in the reductive cycles that lead to the huge variation among naturally occurring polyketides (Fig. 1).

Over the past decade, cloning and sequence analysis of microbial PKS genes have led to the identification of at least three architecturally different types of PKSs (Refs 2,3). Complex or 'modular' PKSs, which catalyze the biosynthesis of macrolides such as erythromycin<sup>4,5</sup>, avermectin<sup>6</sup> and rapamycin<sup>7</sup>, are assemblies of large multifunctional proteins [designated 6-deoxyerythronolide B synthase (DEBS 1, 2 and 3); Fig. 2] that carry a distinct active site for every enzyme-catalyzed step in carbon-chain assembly and modification.

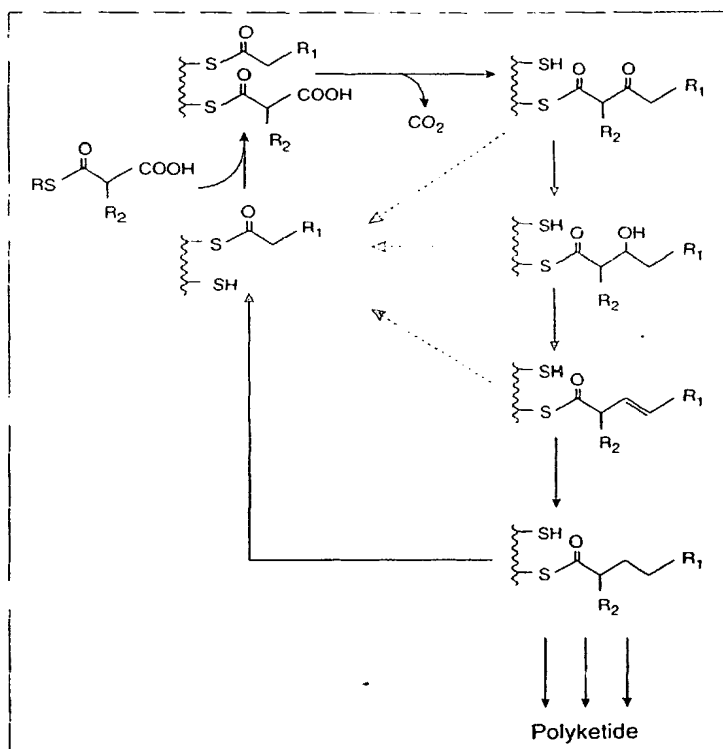


Figure 1

The catalytic cycle of a polyketide synthase (PKS). The growing polyketide chain and each extender unit are attached to the PKS (shown by a wavy line) as thioesters. In the biosynthetic scheme shown here, different PKSs can display variability with regard to the length of the polyketide chain, the choice of monomer incorporated at each step, the degree of reduction of each  $\beta$ -keto group and the stereochemistry at each chiral center. For example, the dashed arrows illustrate how the degree of  $\beta$ -keto-reduction can vary at any given carbonyl. PKSs control the chain length of the product by releasing it after a defined number of condensation and reduction cycles.

C. Khosla (*ckhosla@chemeng.stanford.edu*) and R. J. X. Zawada are at the Department of Chemical Engineering, Stanford University, Stanford, CA 94305-5025, USA.

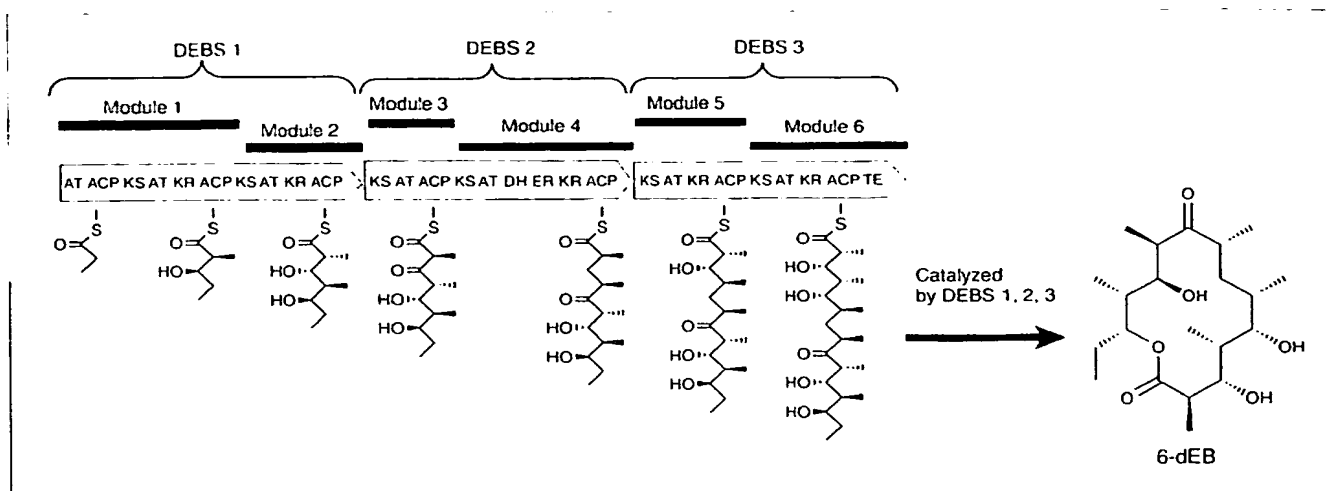


Figure 2

Synthetic pathway catalyzed by 6-deoxyerythronolide B (6-dEB) synthases 1, 2 and 3 (DEBS 1, 2 and 3). The DEBS polyketide synthase (PKS) assembly includes the three multifunctional proteins DEBS 1, 2 and 3, each of which comprises two modules. Each module contains a full complement of sites required for one condensation, an acyl transferase (AT) [a β-ketoacyl carrier protein synthase (KS) and an acyl carrier protein (ACP)], along with a subset of reductive active sites [β-ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER)]. The extent of reduction a particular condensation undergoes depends upon which reducing sites the corresponding module includes. After the final condensation and reduction, the thioesterase (TE) cleaves the polyketide chain from the enzyme.

Active sites are clustered into modules, with each module comprising a full complement of sites required for one condensation and associated reduction cycle. By contrast, the PKSs responsible for the biosynthesis of bacterial aromatic polyketides such as actinorhodin (*act*; Ref. 8), granaticin (*gra*; Ref. 9) and tetracenomycin (*tcm*; Ref. 10) are made up of fewer active sites, each of which is individually encoded as a distinct polypeptide. Because a given type of active site is represented only once within such a PKS gene set (Fig. 3), it has been proposed that some active sites are used iteratively in the biosynthesis of a polyketide molecule.

Fungal PKSs, such as the 6-methylsalicylic acid synthase<sup>11</sup>, are interesting hybrids between the above two types of PKSs: a single gene encodes a multi-domain polypeptide that contains all the active sites required for the biosynthesis of the polyketide. Thus, whereas their domain organization resembles that of modular PKSs, the active sites appear to be used iteratively, as in the case of bacterial aromatic PKSs.

Inspired by the structural complexity and pharmaceutical relevance of their products, the molecular-recognition features of PKSs have been the target of intensive manipulation and analysis via genetic engineering. Recent studies have provided interesting insights into the enzymatic basis for pathway specificity, and have studied the synthesis of increasing numbers of 'unnatural' natural products.

#### Combinatorial biosynthesis

Combinatorial biosynthesis involves genetically engineering biosynthetic pathways in such a way that they may be combinatorially reconstructed to produce libraries of novel small molecules that are appropriate for use in screening for new drugs. Since the first report of the use of genetic engineering to produce

'hybrid' polyketides<sup>12</sup>, the idea of recombining naturally occurring biosynthetic pathways to create new organic molecules has received much attention. However, until recently, the literature contained only a few successful examples that illustrated this principle<sup>5,13-15</sup>. One reason for the limited exploitation of this technique was that the construction of recombinant microorganisms with the ability to produce new molecules relied upon the use of relatively inefficient and time-consuming methods such as genome replacement and/or complementation of pathway mutants. The situation changed with the recent development of a novel host-vector strategy in *Streptomyces coelicolor*<sup>16</sup>, which can be used to express gene clusters for entire PKS pathways (wild type or recombinant) in a manner that is amenable to well-established high-frequency mutagenesis technologies<sup>16-18</sup>. This has made possible the generation of large-scale clonal libraries of recombinant microorganisms, each containing a unique set of biosynthetic genes. In theory, the number of polyketide products that can be obtained using this approach is given by the formula  $R^n$ , where  $R$  is the number of different genes that can be used in each construct and  $n$  is the number of different allelic forms of each gene (i.e. the number of source organisms) that can be obtained. Thus, if two polyketide pathways, each containing four homologous polyketide biosynthesis genes with different molecular-recognition features, are combined in all of the possible ways, 16 different polyketides could be synthesized. Similarly, four pathways with four genes could be combined to generate 256 unique polyketides—hence the term 'combinatorial biosynthesis'. The resulting polyketides can then either be further modified or directly screened to identify promising drug leads.

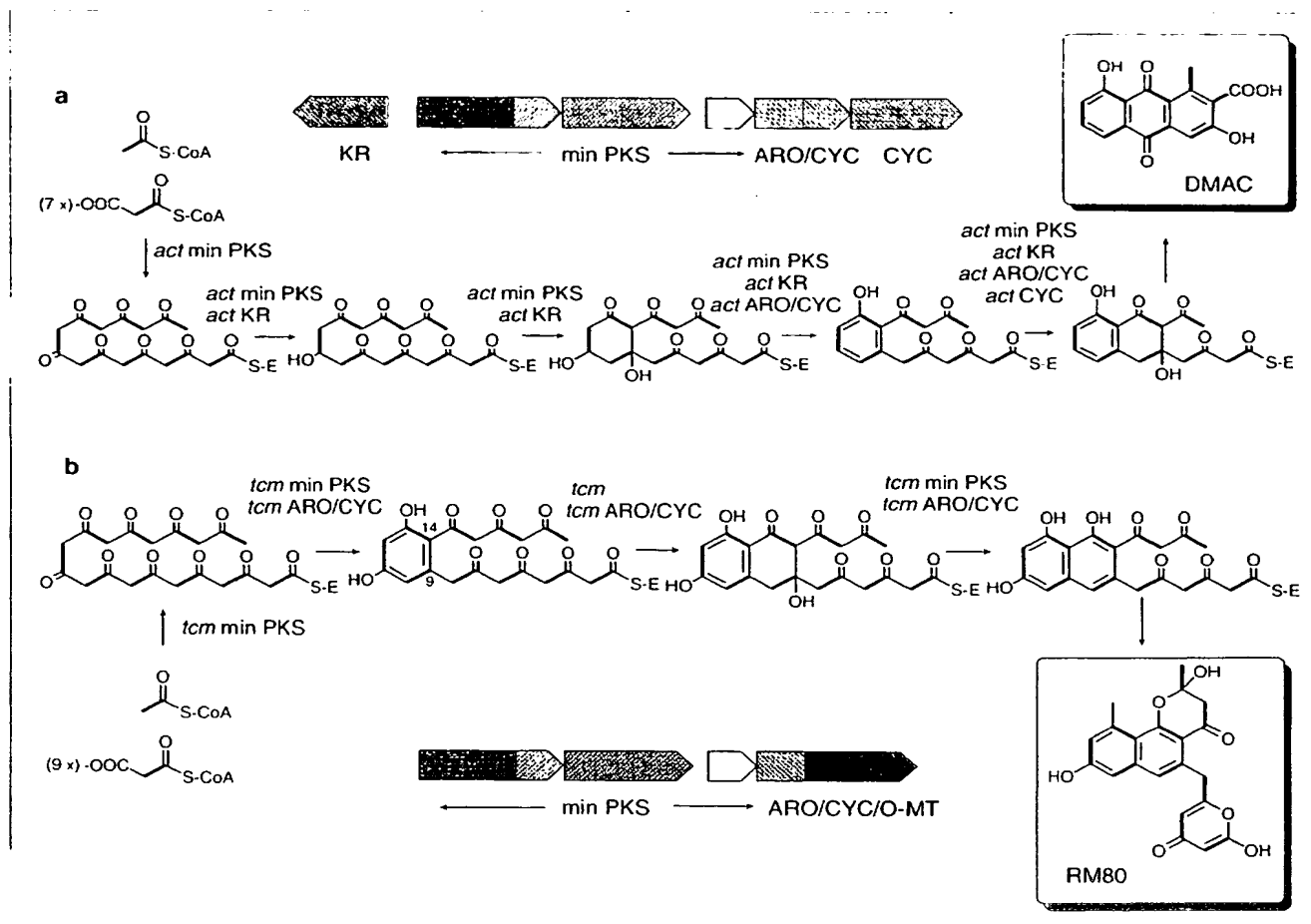


Figure 3

Gene clusters and biosynthetic pathways for two aromatic polyketide synthases (PKSs). (a) Gene cluster and biosynthetic pathway for the actinorhodin (*act*) PKS. The *act* minimal (min) PKS (three gene products) is involved in condensing one acetyl and seven malonyl thioesters to form an octaketide backbone, which is then modified into DMAC by the *act* min PKS in conjunction with the *act* ketoreductase (KR), *act* aromatase/cyclase (ARO/CYC) and *act* cyclase (CYC). (b) Gene cluster and biosynthetic pathway for the tetracenomycin (*tcm*) PKS. The *tcm* min PKS catalyzes the formation of a 20-carbon backbone, which is modified into RM80 by the *tcm* min PKS and *tcm* ARO/CYC.

Of central relevance to the design of a combinatorial polyketide library is an understanding of the genetic basis for control of the catalytic functions and associated molecular-recognition features within a PKS, and the extent to which these properties can be independently manipulated. Thus, starting from a given PKS as a scaffold, the size and diversity of the resulting polyketide library depend on the number of ways that an existing function can be deleted, a new function can be added or the recognition features associated with a given function can be altered without the remainder of the catalytic cycle being affected. These questions have been intensively explored in the context of the bacterial aromatic PKSs.

#### Libraries derived from aromatic PKSs

Figure 3 illustrates the biosynthetic pathways and gene clusters for two closely related systems, the *act* and *tcm* PKSs. Notwithstanding the overall similarities, key differences can be observed. For example, whereas

the *act* PKS catalyzes the formation of an initial C-7–C-12 carbocycle<sup>19</sup> (an aldol condensation between the C-7 carbonyl and C-12 methylene, which results in a six carbon ring), the *tcm* PKS catalyzes a C-9–C-14 cyclization (an aldol condensation between the C-9 carbonyl and C-14 methylene) in its nascent polyketide intermediate<sup>20,21</sup>. It was shown that the latter property is controlled by the *tcm* aromatase/cyclase ARO/CYC. Based on this finding, it was possible to design a hybrid PKS, composed of the *act* minimal PKS (which catalyzes biosynthesis of an unreduced octaketide backbone) and *tcm* ARO/CYC; the result was a novel product (RM77: 1) that possessed the relevant features of both the natural products<sup>20,21</sup> (Fig. 4a). Similarly, the *act* ketoreductase (KR), which controls regiospecific reduction of the octaketide backbone at C 9 (a feature that is absent in the overall catalytic cycle of the *tcm* PKS), can be combined with the *tcm* minimal PKS to produce decaketides RM20b (2) and RM20 (3) that undergo C-9 reduction (Fig. 4b; Refs 16,22).

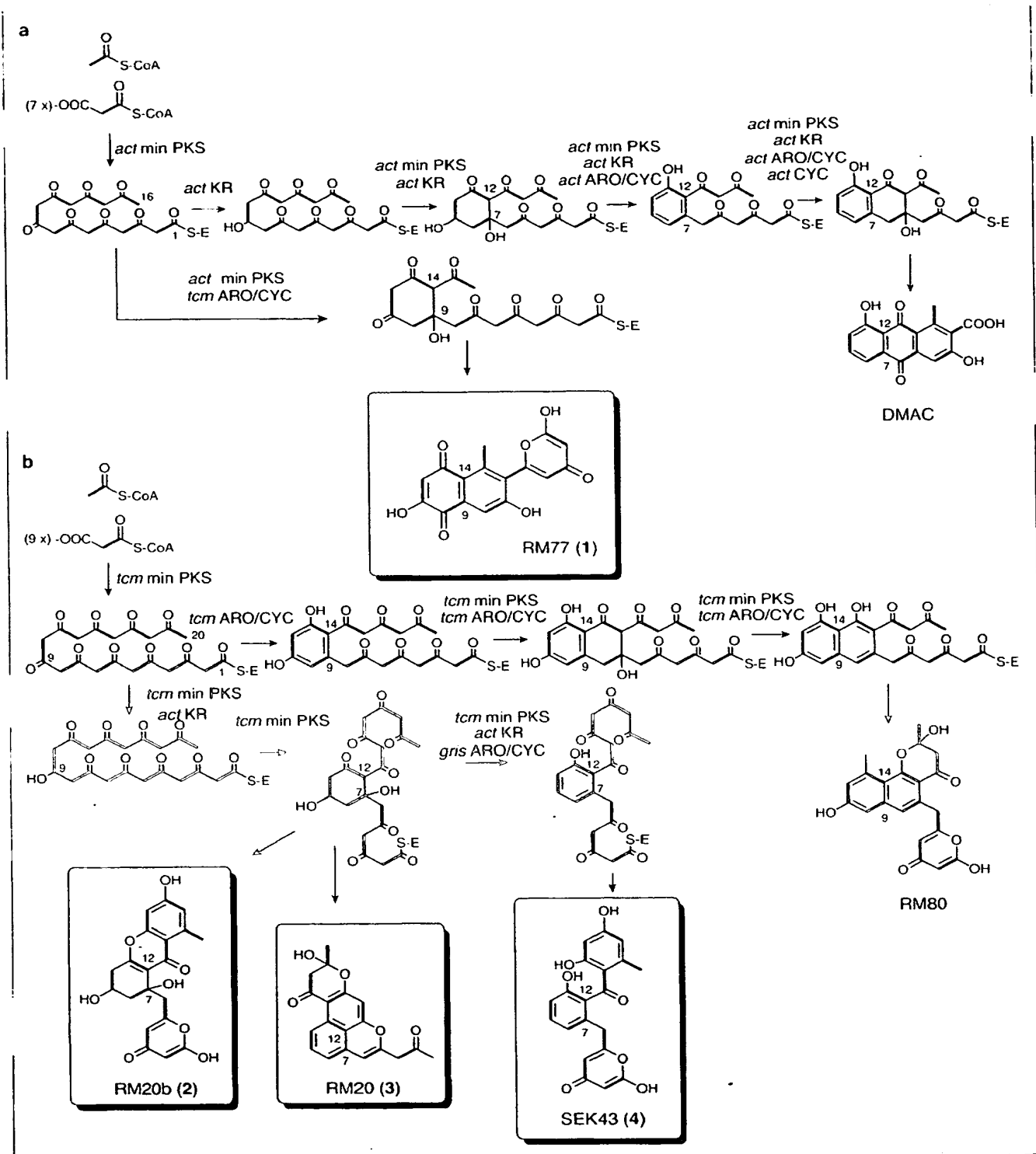


Figure 4

Combinatorial biosynthesis with aromatic polyketide synthases (PKSs). (a) Combining the actinorhodin minimal (*act min*) PKS and tetracenomycin (*tcm*) aromatase/cyclase (ARO/CYC). The *act min* PKS synthesizes a 16-carbon polyketide chain on which the *tcm* ARO/CYC catalyzes a C-9-C-14 cyclization (condensation between C-9 carbonyl and C-14 methylene) leading to the formation of RM77 (1). The native *act* PKS catalyzes a C-7-C-12 cyclization in the synthesis of DMAC. (b) Combining *tcm min* PKS, *act* KR and griseusin (*gns*) ARO/CYC. Adding the *act* KR to the *tcm min* PKS results in the formation of a 20-carbon polyketide in which the C-9 carbonyl is reduced, and there is a carbocycle between C-7 and C-12. RM20b (2) and RM20 (3) are the resulting products. Including the *gns* ARO/CYC with the *tcm min* PKS and *act* KR results in a product SEK43 (4), in which the initial ring is aromatized. The native *tcm* PKS does not reduce the C-9 carbonyl, but catalyzes a C-9-C-14 cyclization in the synthesis of RM80.

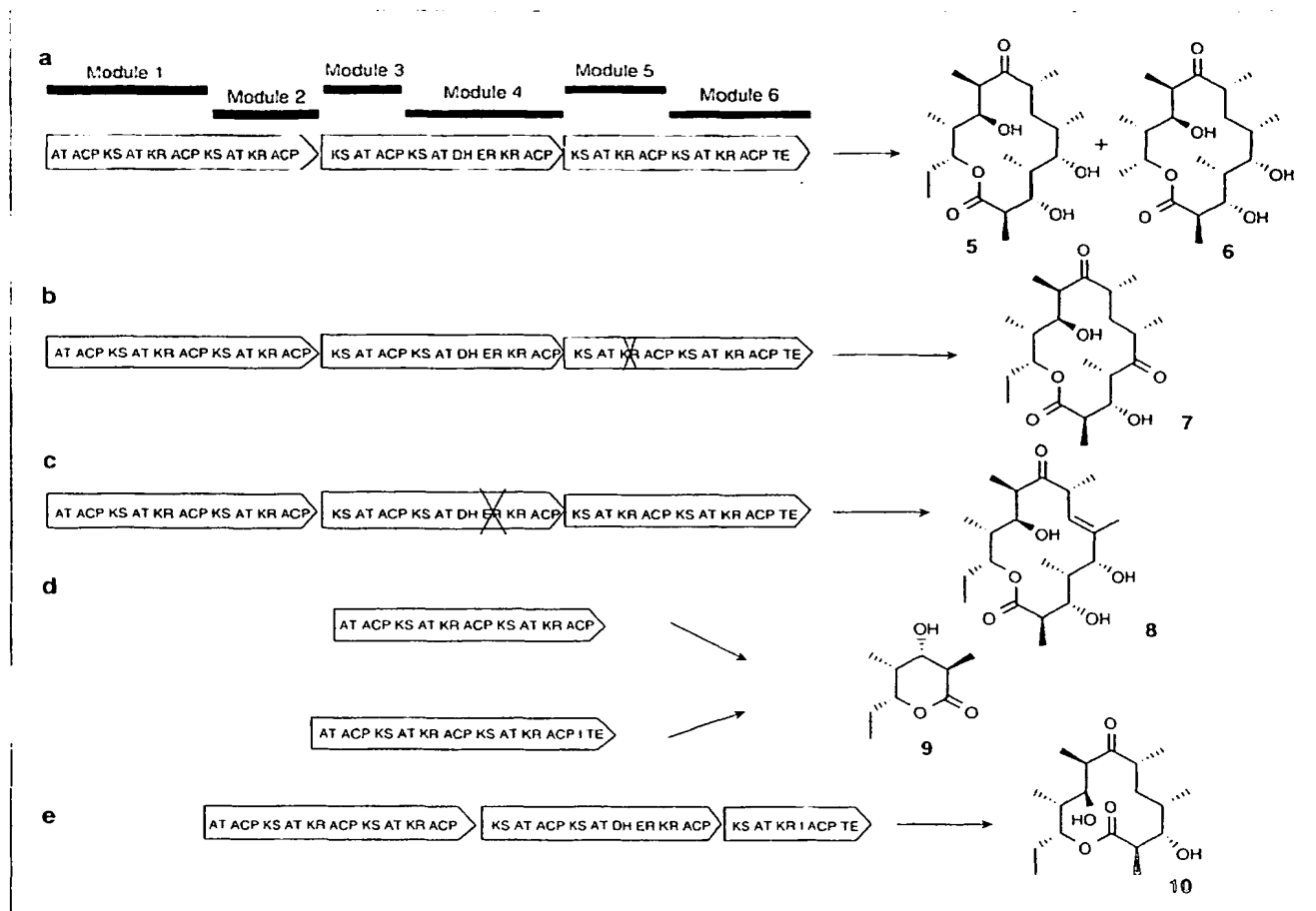


Figure 5

Various genetic manipulations of the 6-deoxyerythronolide B synthase (DEBS) genes. (a) Expression of the complete DEBS gene cluster results in the production of 6-deoxyerythronolide B (6-DEB; **5**), which is derived from a propionate primer unit, and 8,8a-deoxyoleandolide (**6**), which is derived from an acetate primer unit. (b) Deletion mutagenesis of the ketoreductase (KR) domain of module 5 results in the product 5-oxo-6-deoxyerythronolide B (**7**), which is not reduced following the fifth condensation. (c) Deletion of the enoylreductase (ER) domain of module 4 leads to the production of a C-6-C-7 anhydro macrolactone (**8**), which is not reduced following dehydration by the dehydratase (DH) domain of module 4. (d) When only a single DEBS protein [DEBS 1, or DEBS 1 with the thioesterase (TE) from DEBS 3 fused to its C-terminal end] is expressed, a triketide lactone (**9**) is generated. (e) Fusing the acyl carrier protein (ACP)-TE di-domain from module 6 downstream from module 5, which effectively eliminates module 6, results in a 12-membered macrolactone (**10**) as opposed to the 14-membered macrolactones (**5** and **6**) that are produced when module 6 is also included. Crosses indicate the deletion of a particular domain and horizontal lines mark the position of a fusion. Abbreviations: AT, acyl transferase; KS,  $\beta$ -ketoacyl carrier protein synthase.

This combinatorial principle can be extended to more than two PKSs. For example, the *ton* minimal PKS, the *act* ketoreductase and the griseusin (*gris*) ARO/CYC (which, like its *act* counterpart, catalyzes aromatization of the first ring of a reduced polyketide, but unlike its *act* counterpart, can recognize a deca-ketide substrate) can productively associate to produce SEK-43 (**4**), a product that exhibits the relevant features of all three parent natural products (Fig. 4b; Ref. 23). Indeed, the concept of combinatorial biosynthesis has already been exploited to generate a sizeable library of 'unnatural' natural products using five naturally occurring aromatic PKS gene clusters as a starting point (reviewed in Ref. 24). Based on current knowledge, it has been estimated that this library could be expanded to include at least a few hundred novel products<sup>24</sup>.

#### Libraries derived from modular PKSs

In contrast to bacterial aromatic PKSs, our current knowledge of the combinatorial potential in modular PKSs is relatively limited. However, the one-to-one correspondence between active sites and product structure (Fig. 2), together with the incredible chemical diversity observed among their natural products, suggests that the potential for genetically manipulating these multi-enzyme systems could be considerably greater than that for bacterial aromatic PKSs (Ref. 24). For example, a wider range of primer units including aliphatic (e.g. acetate, propionate, butyrate, isovalerate), aromatic (benzoate, aminohydroxybenzoate) and alicyclic (cyclohexanoate) monomers are found in various macrocyclic polyketides. Studies have shown that modular PKSs have relaxed specificity for their primer units. For example, heterologous expression of the

entire DEBS gene cluster in a heterologous host resulted in production of not only 6-deoxyerythronolide B [6-dEB (5)] but also 8.8a-deoxyoleandolide (6), an analog derived from an acetate primer unit (Fig. 5a; Ref. 17).

Modular PKSs also exhibit considerable variety concerning the choice of extender units in each condensation cycle, although it remains to be seen to what extent this property can be manipulated. Moreover, the degree of  $\beta$ -ketoreduction following a condensation reaction can be altered by genetic manipulation. Thus, deletion mutagenesis in the ketoreductase domain of module 5 of the DEBS genes in *Saccharopolyspora erythraea* (an erythromycin producer) resulted in formation of the predicted 5-oxo-6-deoxyerythronolide B product (7), indicating that active sites within module 6 can recognize and process the 12-carbon oxo-intermediate (Fig. 5b; Ref. 5). Similarly, directed mutagenesis of a conserved motif in the enoylreductase domain of module 4 in *S. erythraea* led to production of the expected C-6-C-7 anhydro-macrolactone (8; Fig. 5c; Ref. 15).

The size and shape of the polyketide product can be varied by designing mutants with the appropriate number of modules. For example, a 'designer' triketide lactone, (2R,3S,4S,5R)-2,4-dimethyl-3,5-dihydroxy-*n*-heptanoic acid  $\delta$  lactone (9; C-9 lactone) produced by DEBS 1 alone<sup>25</sup>, as well as by the fusion protein DEBS 1+TE, which contains the thioesterase (TE) domain from DEBS 3 fused to the C-terminal end of DEBS 1 (Fig. 5d; Refs 26,27). Similarly, a new 12-membered macrolactone, (8S,9S)-8,9-dihydro-8-methyl-9-hydroxy-10-deoxymethynolide (10), was produced by fusing the ACP-TE di-domain of module 6 downstream of the KR domain of module 5 (Fig. 5e; Ref. 26). In addition to providing insights into the regiospecificity of cyclization by the TE domain, this result also illustrated the feasibility of constructing PKSs containing hybrid modules. Further manipulations along these lines in DEBS and other modular PKS gene clusters will undoubtedly yield valuable information regarding the scope and limitations of combinatorial biosynthesis using these remarkable multi-enzyme assemblies.

## Conclusions

As a family, PKSs provide a fundamentally interesting and medically important example of a remarkable strategy used by nature to generate molecular diversity via a building-block assembly of complex organic molecules. In contrast to multi-enzyme systems such as those responsible for replication, transcription and translation, where diversity arises from the structure of a modular template, PKSs represent a case in which the protein catalysts are themselves modularized. This inherent modularity at the enzymatic level can be exploited via combinatorial biosynthesis.

Although the field is still in its infancy, based on available information, there is every reason to look forward to a future when libraries of complex natural

product-like molecules are generated in the form of spatially segregated bacterial clones on a petri dish. In an era where the inherent costs and redundancies existing in conventional natural product-based drug-discovery programs are increasingly being called into question, combinatorial biosynthesis offers a unique opportunity for drug discovery and subsequent process development. It is already becoming increasingly clear that the domains and subunits from naturally occurring PKSs can be genetically recombined to produce novel molecules. As our understanding of the relationship between structure, mechanisms and molecular-recognition features of these remarkable multi-enzyme assemblies improves, it may be possible to redesign naturally occurring domains and subunits to exhibit properties not ordinarily seen in natural systems.

Finally, although beyond the scope of this review, the possibility of using PKSs as 'one-pot' catalysts for *in vitro* combinatorial chemistry is also becoming increasingly feasible with the advent of efficient cell-free systems capable of enzymatic synthesis of polyketides<sup>28-33</sup>.

## Acknowledgements

Research in the authors' laboratory was supported in part by grants from the National Institutes of Health (CA 66736-01), the National Science Foundation (MCB9417419), a National Science Foundation Young Investigator Award and a David and Lucile Packard Fellowship for Science and Engineering.

## References

- O'Hagan, D. (1991) *The Polyketide Metabolites*, Ellis Horwood.
- Katz, L. and Donadio, S. (1993) *Annu. Rev. Microbiol.* 47, 875-912.
- Hutchinson, C. R. and Fujii, I. (1995) *Annu. Rev. Microbiol.* 49, 201-238.
- Cortes, J., Haydock, S. F., Roberts, G. A., Beville, D. J. and Leadlay, P. F. (1990) *Nature* 348, 176-178.
- Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. and Katz, L. (1991) *Science* 252, 675-679.
- MacNeil, D. J. et al. (1992) *Gene* 115, 119-125.
- Schweke, T. et al. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 7839-7843.
- Fernández-Moreno, M. A., Martínez, E., Boto, L., Hopwood, D. A. and Malparada, F. (1992) *J. Biol. Chem.* 267, 19278-19290.
- Sherman, D. H. et al. (1989) *EMBO J.* 8, 2717-2725.
- Bibb, M. J., Biró, S., Motamedi, H., Collins, J. F. and Hutchinson, C. R. (1989) *EMBO J.* 8, 2727-2736.
- Beck, J., Ripka, S., Siegner, A., Schultz, E. and Schweizer, E. (1990) *Eur. J. Biochem.* 192, 487-498.
- Hopwood, D. A. et al. (1985) *Nature* 314, 642-644.
- Epp, J. K. et al. (1989) *Gene* 85, 293-301.
- Bartel, P. L. et al. (1990) *J. Bacteriol.* 172, 4816-4826.
- Donadio, S., McAlpine, J. B., Sheldon, P. J., Jackson, M. and Katz, L. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 7119-7123.
- McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. and Khosla, C. (1993) *Science* 262, 1546-1550.
- Kao, C. M., Katz, L. and Khosla, C. (1994) *Science* 265, 509-512.
- Bedford, D. J., Schweizer, E., Hopwood, D. A. and Khosla, C. (1995) *J. Bacteriol.* 177, 4544-4548.
- McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. and Khosla, C. (1994) *J. Am. Chem. Soc.* 116, 10855-10859.
- McDaniel, R., Hutchinson, C. R. and Khosla, C. (1995) *J. Am. Chem. Soc.* 117, 6805-6810.

- 21 Shen, B., Summers, R. G., Wendt-Pienkowska, E. and Hutchinson, C. R. (1995) *J. Am. Chem. Soc.* 117, 6811-6821
- 22 Fu, H., McDaniel, R., Hopwood, D. A. and Khosla, C. (1994) *Biochemistry* 33, 9321-9326
- 23 McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. and Khosla, C. (1995) *Nature* 375, 549-554
- 24 Tsou, C. J. and Khosla, C. (1995) *Chem. Biol.* 2, 355-362
- 25 Kao, C. M., Luo, G., Katz, L., Cane, D. E. and Khosla, C. (1994) *J. Am. Chem. Soc.* 116, 11612-11613
- 26 Kao, C. M., Luo, G., Katz, L., Cane, D. E. and Khosla, C. (1995) *J. Am. Chem. Soc.* 117, 9105-9106
- 27 Cortes, J. *et al.* (1995) *Science* 268, 1487-1489
- 28 Shen, B. and Hutchinson, C. R. (1993) *Science* 262, 1535-1540
- 29 Pieper, R., Luo, G., Cane, D. E. and Khosla, C. (1995) *Nature* 378, 263-266
- 30 Pieper, R., Luo, G., Cane, D. E. and Khosla, C. (1995) *J. Am. Chem. Soc.* 117, 11373-11374
- 31 Wiesmann, K. E. H. *et al.* (1995) *Chem. Biol.* 2, 583-589
- 32 Luo, G., Pieper, R., Khosla, C. and Cane, D. E. *Bioorg. Med. Chem.* 4, 995-999
- 33 Carreras, C. W., Pieper, R. and Khosla, C. *J. Am. Chem. Soc.* (in press)

*The ability to reproduce reliably the hematopoietic differentiation and expansion process in vitro would greatly facilitate the development of clinical therapeutic treatments based on blood products and cell transplantation. However, the technical challenges to developing cost-effective productive hematopoietic cell culture systems are significant: the complexity of the process requires an improved understanding of the biology, and the engineering of culture systems that provide the appropriate biochemical and physical environment, and the necessary control over culture parameters. In this first part of a two-part article, E. T. Papoutsakis and colleagues present the cell-culture considerations; in the second part (to be published in the October issue of TIBTECH), they discuss the therapeutic applications that will result from such improved culture systems.*

## Hematopoietic cell culture therapies (Part I): cell culture considerations

Todd A. McAdams, William M. Miller and E. Terry Papoutsakis

Hematopoietic cell culture, or *ex vivo* expansion of hematopoietic cells, is a rapidly growing area of tissue engineering with many potential applications in bone-marrow transplantation, gene therapy and the production of blood products. Hematopoietic cultures are considerably more complex than established animal cell culture technologies owing to the presence of many cell types at various differentiation stages, and stringent medium and growth factor requirements. Culture parameters, such as pH, oxygen tension, seeding density and feeding schedules, significantly affect the proliferation and differentiation of hematopoietic cells. A number of bioreactor systems are currently under development.

Hematopoiesis is the process of generating mature blood cells (Box 1), which are produced at a rate of 400 billion per day in the average human<sup>1</sup>. Since the first successful human hematopoietic cultures, advances in cytokine discovery, flow cytometry and cell selection technologies have greatly expanded the

tools available for understanding hematopoietic differentiation, and made possible the controlled *ex vivo* expansion of hematopoietic cells. Nevertheless, hematopoietic cultures remain among the most challenging culture systems because of the intrinsically heterogeneous and highly variable nature of the cultured cells. Heterogeneity and variability derive both from the cell source (patient to patient variability) and the complexity of hematopoietic differentiation (Fig. 1). The hematopoietic stem cell can develop into multiple lineages, with a continuum of differentiation

*T. A. McAdams, W. M. Miller and E. T. Papoutsakis (e-paps@nwu.edu) are at the Department of Chemical Engineering, Northwestern University, Evanston, IL 60208-3120, U.S.A. Address reprint requests and correspondence to E. T. P.*

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**